

#5

OIKE

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/755,830

DATE: 08/17/2001
 TIME: 12:02:58

Input Set : A:\10287-067001.TXT
 Output Set: N:\CRF3\08162001\I755830.raw

4 <110> APPLICANT: Georgopoulos, Katia
 6 <120> TITLE OF INVENTION: IKAROS REGULATORY ELEMENTS AND USES
 7 THEREOF
 9 <130> FILE REFERENCE: 10287-067001
 11 <140> CURRENT APPLICATION NUMBER: US 09/755,830
 12 <141> CURRENT FILING DATE: 2001-01-05
 14 <150> PRIOR APPLICATION NUMBER: US 08/283,300
 15 <151> PRIOR FILING DATE: 1994-07-29
 17 <150> PRIOR APPLICATION NUMBER: US 08/238,212
 18 <151> PRIOR FILING DATE: 1994-05-02
 20 <150> PRIOR APPLICATION NUMBER: US 08/121,438
 21 <151> PRIOR FILING DATE: 1993-09-14
 23 <150> PRIOR APPLICATION NUMBER: US 07/946,233
 24 <151> PRIOR FILING DATE: 1992-09-14
 26 <160> NUMBER OF SEQ ID NOS: 43
 28 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 30 <210> SEQ ID NO: 1
 31 <211> LENGTH: 1788
 32 <212> TYPE: DNA
 33 <213> ORGANISM: Mus musculus
 35 <220> FEATURE:
 36 <221> NAME/KEY: CDS
 37 <222> LOCATION: (223)...(1515)
 38 <223> OTHER INFORMATION: mik-2
 40 <400> SEQUENCE: 1
 41 aattcgttct accttctctg aaccccgagt gtgtgtcaag gccggactgg gagcttgggg 60
 42 gaagaggaag aggaagagga atctgcggct catccaggga tcagggtcct tcccaagtgg 120
 43 ccactcagag gggactcaga gcaagtctag atttgtgtgg cagagagaga cagctctcgt 180
 44 ttggccttgg ggaggcacaa gtctgttgat aacctgaaga ca atg gat gtc gat 234
 45 Met Asp Val Asp
 46 1
 48 gag ggt caa gac atg tcc caa gtt tca gga aag gag agc ccc cca gtc 282
 49 Glu Gly Gln Asp Met Ser Gln Val Ser Gly Lys Glu Ser Pro Pro Val
 50 5 10 15 20
 52 agt gac act cca gat gaa ggg gat gag ccc atg cct gtc cct gag gac 330
 53 Ser Asp Thr Pro Asp Glu Gly Asp Glu Pro Met Pro Val Pro Glu Asp
 54 25 30 35
 56 ctg tcc act acc tct gga gca cag cag aac tcc aag agt gat cga ggc 378
 57 Leu Ser Thr Thr Ser Gly Ala Gln Gln Asn Ser Lys Ser Asp Arg Gly
 58 40 45 50
 60 atg ggt gaa cgg cct ttc cag tgc aac cag tct ggg gcc tcc ttt acc 426
 61 Met Gly Glu Arg Pro Phe Gln Cys Asn Gln Ser Gly Ala Ser Phe Thr
 62 55 60 65
 64 cag aaa ggc aac ctc ctg cgg cac atc aag ctg cac tcg ggt gag aag 474
 65 Gln Lys Gly Asn Leu Leu Arg His Ile Lys Leu His Ser Gly Glu Lys
 66 70 75 80
 68 ccc ttc aaa tgc cat ctt tgc aac tat gcc tgc cgc cgg agg gac gcc 522

ENTERED

p.5

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/755,830

DATE: 08/17/2001

TIME: 12:02:58

Input Set : A:\10287-067001.TXT

Output Set: N:\CRF3\08162001\I755830.raw

69	Pro	Phe	Lys	Cys	His	Leu	Cys	Asn	Tyr	Ala	Cys	Arg	Arg	Arg	Asp	Ala	
70	85					90					95					100	
72	ctc	acc	ggc	cac	ctg	agg	acg	cac	tcc	gtt	ggt	aag	cct	cac	aaa	tgt	570
73	Leu	Thr	Gly	His	Leu	Arg	Thr	His	Ser	Val	Gly	Lys	Pro	His	Lys	Cys	
74					105					110					115		
76	gga	tat	tgt	ggc	cgg	agc	tat	aaa	cag	cga	agc	tct	tta	gag	gag	cat	618
77	Gly	Tyr	Cys	Gly	Arg	Ser	Tyr	Lys	Gln	Arg	Ser	Ser	Leu	Glu	Glu	His	
78				120					125					130			
80	aaa	gag	cga	tgc	cac	aac	tac	ttg	gaa	agc	atg	ggc	ctt	ccg	ggc	gtg	666
81	Lys	Glu	Arg	Cys	His	Asn	Tyr	Leu	Glu	Ser	Met	Gly	Leu	Pro	Gly	Val	
82			135					140					145				
84	tgc	cca	gtc	att	aag	gaa	gaa	act	aac	cac	aac	gag	atg	gca	gaa	gac	714
85	Cys	Pro	Val	Ile	Lys	Glu	Glu	Thr	Asn	His	Asn	Glu	Met	Ala	Glu	Asp	
86		150					155					160					
88	ctg	tgc	aag	ata	gga	gca	gag	agg	tcc	ctt	gtc	ctg	gac	agg	ctg	gca	762
89	Leu	Cys	Lys	Ile	Gly	Ala	Glu	Arg	Ser	Leu	Val	Leu	Asp	Arg	Leu	Ala	
90	165				170					175					180		
92	agc	aat	gtc	gcc	aaa	cgt	aag	agc	tct	atg	cct	cag	aaa	ttt	ctt	gga	810
93	Ser	Asn	Val	Ala	Lys	Arg	Lys	Ser	Ser	Met	Pro	Gln	Lys	Phe	Leu	Gly	
94				185					190					195			
96	gac	aag	tgc	ctg	tca	gac	atg	ccc	tat	gac	agt	gcc	aac	tat	gag	aag	858
97	Asp	Lys	Cys	Leu	Ser	Asp	Met	Pro	Tyr	Asp	Ser	Ala	Asn	Tyr	Glu	Lys	
98			200					205					210				
100	gag	gat	atg	atg	aca	tcc	cac	gtg	atg	gac	cag	gcc	atc	aac	aat	gcc	906
101	Glu	Asp	Met	Met	Thr	Ser	His	Val	Met	Asp	Gln	Ala	Ile	Asn	Asn	Ala	
102			215					220					225				
104	atc	aac	tac	ctg	ggg	gct	gag	tcc	ctg	cgc	cca	ttg	gtg	cag	aca	ccc	954
105	Ile	Asn	Tyr	Leu	Gly	Ala	Glu	Ser	Leu	Arg	Pro	Leu	Val	Gln	Thr	Pro	
106		230				235					240						
108	ccc	ggt	agc	tcc	gag	gtg	gtg	cca	gtc	atc	agc	tcc	atg	tac	cag	ctg	1002
109	Pro	Gly	Ser	Ser	Glu	Val	Val	Pro	Val	Ile	Ser	Ser	Met	Tyr	Gln	Leu	
110	245				250				255						260		
112	cac	aag	ccc	ccc	tca	gat	ggc	ccc	cca	cgg	tcc	aac	cat	tca	gca	cag	1050
113	His	Lys	Pro	Pro	Ser	Asp	Gly	Pro	Pro	Arg	Ser	Asn	His	Ser	Ala	Gln	
114				265				270					275				
116	gac	gcc	gtg	gat	aac	ttg	ctg	ctg	ctg	tcc	aag	gcc	aag	tct	gtg	tca	1098
117	Asp	Ala	Val	Asp	Asn	Leu	Leu	Leu	Leu	Ser	Lys	Ala	Lys	Ser	Val	Ser	
118			280					285					290				
120	tcg	gag	cga	gag	gcc	tcc	ccg	agc	aac	agc	tgc	caa	gac	tcc	aca	gat	1146
121	Ser	Glu	Arg	Glu	Ala	Ser	Pro	Ser	Asn	Ser	Cys	Gln	Asp	Ser	Thr	Asp	
122			295					300					305				
124	aca	gag	agc	aac	gcg	gag	gaa	cag	cgc	agc	ggc	ctt	atc	tac	cta	acc	1194
125	Thr	Glu	Ser	Asn	Ala	Glu	Glu	Gln	Arg	Ser	Gly	Leu	Ile	Tyr	Leu	Thr	
126		310				315						320					
128	aac	cac	atc	aac	ccg	cat	gca	cgc	aat	ggg	ctg	gct	ctc	aag	gag	gag	1242
129	Asn	His	Ile	Asn	Pro	His	Ala	Arg	Asn	Gly	Leu	Ala	Leu	Lys	Glu	Glu	
130	325				330					335					340		
132	cag	cgc	gcc	tac	gag	gtg	ctg	agg	gcg	gcc	tca	gag	aac	tcg	cag	gat	1290
133	Gln	Arg	Ala	Tyr	Glu	Val	Leu	Arg	Ala	Ala	Ser	Glu	Asn	Ser	Gln	Asp	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/755,830

DATE: 08/17/2001

TIME: 12:02:58

Input Set : A:\10287-067001.TXT

Output Set: N:\CRF3\08162001\I755830.raw

```

134          345          350          355
136 gcc ttc cgt gtg gtc agc acg agt ggc gag cag ctg aag gtg tac aag      1338
137 Ala Phe Arg Val Val Ser Thr Ser Gly Glu Gln Leu Lys Val Tyr Lys
138          360          365          370
140 tgc gaa cac tgc cgc gtg ctc ttc ctg gat cac gtc atg tat acc att      1386
141 Cys Glu His Cys Arg Val Leu Phe Leu Asp His Val Met Tyr Thr Ile
142          375          380          385
144 cac atg ggc tgc cat ggc tgc cat ggc ttt cgg gat ccc ttt gag tgt      1434
145 His Met Gly Cys His Gly Cys His Gly Phe Arg Asp Pro Phe Glu Cys
146          390          395          400
148 aac atg tgt ggt tat cac agc cag gac agg tac gag ttc tca tcc cat      1482
149 Asn Met Cys Gly Tyr His Ser Gln Asp Arg Tyr Glu Phe Ser Ser His
150 405          410          415          420
152 atc acg cgg ggg gag cat cgt tac cac ctg agc taaacccagc caggccccac      1535
153 Ile Thr Arg Gly Glu His Arg Tyr His Leu Ser
154          425          430
156 tgaagcacaa agatagctgg ttatgcctcc ttcccggcag ctggacccac agcggacaat      1595
157 gtgggagtggt atttgaggc agcatttggt cttttatggt ggttggttggt cgtttcattt      1655
158 gcgttggaag ataagttttt aatgttagtg acaggattgc attgcatcag caacattcac      1715
159 aacatccatc cttctagcca gttttgttca ctggtagctg aggtttcccg gatatgtggt      1775
160 ttcctaacac tct      1788
162 <210> SEQ ID NO: 2
163 <211> LENGTH: 1386
164 <212> TYPE: DNA
165 <213> ORGANISM: Homo sapiens
167 <220> FEATURE:
168 <221> NAME/KEY: CDS
169 <222> LOCATION: (1)...(1383)
170 <223> OTHER INFORMATION: hik-1
172 <400> SEQUENCE: 2
173 aat gtt aaa gta gag act cag agt gat gaa gag aat ggg cgt gcc tgt      48
174 Asn Val Lys Val Glu Thr Gln Ser Asp Glu Glu Asn Gly Arg Ala Cys
175 1          5          10          15
177 gaa atg aat ggg gaa gaa tgt gcg gag gat tta cga atg ctt gat gcc      96
178 Glu Met Asn Gly Glu Glu Cys Ala Glu Asp Leu Arg Met Leu Asp Ala
179          20          25          30
181 tcg gga gag aaa atg aat ggc tcc cac agg gac caa ggc agc tcg gct      144
182 Ser Gly Glu Lys Met Asn Gly Ser His Arg Asp Gln Gly Ser Ser Ala
183          35          40          45
185 ttg tcg gga gtt gga ggc att cga ctt cct aac gga aaa cta aag tgt      192
186 Leu Ser Gly Val Gly Gly Ile Arg Leu Pro Asn Gly Lys Leu Lys Cys
187          50          55          60
189 gat atc tgt ggg atc att tgc atc ggg ccc aat gtg ctc atg gtt cac      240
190 Asp Ile Cys Gly Ile Ile Cys Ile Gly Pro Asn Val Leu Met Val His
191 65          70          75          80
193 aaa aga agc cac act gga gaa cgg ccc ttc cag tgc aat cag tgc ggg      288
194 Lys Arg Ser His Thr Gly Glu Arg Pro Phe Gln Cys Asn Gln Cys Gly
195          85          90          95
197 gcc tca ttc acc cag aag ggc aac ctg ctc cgg cac atc aag ctg cat      336

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/755,830

DATE: 08/17/2001

TIME: 12:02:58

Input Set : A:\10287-067001.TXT

Output Set: N:\CRF3\08162001\I755830.raw

198	Ala	Ser	Phe	Thr	Gln	Lys	Gly	Asn	Leu	Leu	Arg	His	Ile	Lys	Leu	His	
199				100				105						110			
201	tcc	ggg	gag	aag	ccc	ttc	aaa	tgc	cac	ctc	tgc	aac	tac	gcc	tgc	cgc	384
202	Ser	Gly	Glu	Lys	Pro	Phe	Lys	Cys	His	Leu	Cys	Asn	Tyr	Ala	Cys	Arg	
203			115					120					125				
205	cgg	agg	gac	gcc	ctc	act	ggc	cac	ctg	agg	acg	cac	tcc	gtt	ggt	aaa	432
206	Arg	Arg	Asp	Ala	Leu	Thr	Gly	His	Leu	Arg	Thr	His	Ser	Val	Gly	Lys	
207		130					135					140					
209	cct	cac	aaa	tgt	gga	tat	tgt	ggc	cga	agc	tat	aaa	cag	cga	acg	tct	480
210	Pro	His	Lys	Cys	Gly	Tyr	Cys	Gly	Arg	Ser	Tyr	Lys	Gln	Arg	Thr	Ser	
211	145					150					155					160	
213	tta	gag	gaa	cat	aaa	gag	cgc	tgc	cac	aac	tac	ttg	gaa	agc	atg	ggc	528
214	Leu	Glu	Glu	His	Lys	Glu	Arg	Cys	His	Asn	Tyr	Leu	Glu	Ser	Met	Gly	
215				165				170							175		
217	ctt	ccg	ggc	aca	ctg	tac	cca	gtc	att	aaa	gaa	gaa	act	aag	cac	agt	576
218	Leu	Pro	Gly	Thr	Leu	Tyr	Pro	Val	Ile	Lys	Glu	Glu	Thr	Lys	His	Ser	
219				180				185						190			
221	gaa	atg	gca	gaa	gac	ctg	tgc	aag	ata	gga	tca	gag	aga	tct	ctc	gtg	624
222	Glu	Met	Ala	Glu	Asp	Leu	Cys	Lys	Ile	Gly	Ser	Glu	Arg	Ser	Leu	Val	
223			195					200					205				
225	ctg	gac	aga	cta	gca	agt	aat	gtc	gcc	aaa	cgt	aag	agc	tct	atg	cct	672
226	Leu	Asp	Arg	Leu	Ala	Ser	Asn	Val	Ala	Lys	Arg	Lys	Ser	Ser	Met	Pro	
227		210					215					220					
229	cag	aaa	ttt	ctt	ggg	gac	aag	ggc	ctg	tcc	gac	acg	ccc	tac	gac	agt	720
230	Gln	Lys	Phe	Leu	Gly	Asp	Lys	Gly	Leu	Ser	Asp	Thr	Pro	Tyr	Asp	Ser	
231	225				230					235					240		
233	gcc	acg	tac	gag	aag	gag	aac	gaa	atg	atg	aag	tcc	cac	gtg	atg	gac	768
234	Ala	Thr	Tyr	Glu	Lys	Glu	Asn	Glu	Met	Met	Lys	Ser	His	Val	Met	Asp	
235				245				250						255			
237	caa	gcc	atc	aac	aac	gcc	atc	aac	tac	ctg	ggg	gcc	gag	tcc	ctg	cgc	816
238	Gln	Ala	Ile	Asn	Asn	Ala	Ile	Asn	Tyr	Leu	Gly	Ala	Glu	Ser	Leu	Arg	
239			260					265					270				
241	ccg	ctg	gtg	cag	acg	ccc	ccg	ggc	ggt	tcc	gag	gtg	gtc	ccg	gtc	atc	864
242	Pro	Leu	Val	Gln	Thr	Pro	Pro	Gly	Gly	Ser	Glu	Val	Val	Pro	Val	Ile	
243			275					280					285				
245	agc	ccg	atg	tac	cag	ctg	cac	agg	cgc	tcg	gag	ggc	acc	ccg	cgc	tcc	912
246	Ser	Pro	Met	Tyr	Gln	Leu	His	Arg	Arg	Ser	Glu	Gly	Thr	Pro	Arg	Ser	
247		290					295					300					
249	aac	cac	tcg	gcc	cag	gac	agc	gcc	gtg	gag	tac	ctg	ctg	ctg	ctc	tcc	960
250	Asn	His	Ser	Ala	Gln	Asp	Ser	Ala	Val	Glu	Tyr	Leu	Leu	Leu	Leu	Ser	
251	305				310					315					320		
253	aag	gcc	aag	ttg	gtg	ccc	tcg	gag	cgc	gag	gcg	tcc	ccg	agc	aac	agc	1008
254	Lys	Ala	Lys	Leu	Val	Pro	Ser	Glu	Arg	Glu	Ala	Ser	Pro	Ser	Asn	Ser	
255				325				330					335				
257	tgc	caa	gac	tcc	acg	gac	acc	gag	agc	aac	aac	gag	gag	cag	cgc	agc	1056
258	Cys	Gln	Asp	Ser	Thr	Asp	Thr	Glu	Ser	Asn	Asn	Glu	Glu	Gln	Arg	Ser	
259			340					345					350				
261	ggt	ctt	atc	tac	ctg	acc	aac	cac	atc	gcc	cga	cgc	gcg	caa	cgc	gtg	1104
262	Gly	Leu	Ile	Tyr	Leu	Thr	Asn	His	Ile	Ala	Arg	Arg	Ala	Gln	Arg	Val	

RAW SEQUENCE LISTING

DATE: 08/17/2001

PATENT APPLICATION: US/09/755,830

TIME: 12:02:58

Input Set : A:\10287-067001.TXT

Output Set: N:\CRF3\08162001\I755830.raw

```

263          355          360          365
265 tcg ctc aag gag gag cac cgc gcc tac gac ctg ctg cgc gcc gcc tcc      1152
266 Ser Leu Lys Glu Glu His Arg Ala Tyr Asp Leu Leu Arg Ala Ala Ser
267          370          375          380
269 gag aac tcg cag gac gcg ctc cgc gtg gtc agc acc agc ggg gag cag      1200
270 Glu Asn Ser Gln Asp Ala Leu Arg Val Val Ser Thr Ser Gly Glu Gln
271 385          390          395          400
273 atg aag gtg tac aag tgc gaa cac tgc cgg gtg ctc ttc ctg gat cac      1248
274 Met Lys Val Tyr Lys Cys Glu His Cys Arg Val Leu Phe Leu Asp His
275          405          410          415
277 gtc atg tac acc atc cac atg ggc tgc cac ggc ttc cgt gat cct ttt      1296
278 Val Met Tyr Thr Ile His Met Gly Cys His Gly Phe Arg Asp Pro Phe
279          420          425          430
281 gag tgc aac atg tgc ggc tac cac agc cag gac cgg tac gag ttc tcg      1344
282 Glu Cys Asn Met Cys Gly Tyr His Ser Gln Asp Arg Tyr Glu Phe Ser
283          435          440          445
285 tcg cac ata acg cga ggg gag cac cgc ttc cac atg agc taa      1386
286 Ser His Ile Thr Arg Gly Glu His Arg Phe His Met Ser
287          450          455          460
290 <210> SEQ ID NO: 3
291 <211> LENGTH: 1296
292 <212> TYPE: DNA
293 <213> ORGANISM: Mus musculus
295 <220> FEATURE:
296 <221> NAME/KEY: CDS
297 <222> LOCATION: (1)...(1296)
298 <223> OTHER INFORMATION: mIk-3
300 <400> SEQUENCE: 3
301 atg gat gtc gat gag ggt caa gac atg tcc caa gtt tca gga aag gag      48
302 Met Asp Val Asp Glu Gly Gln Asp Met Ser Gln Val Ser Gly Lys Glu
303 1          5          10          15
305 agc ccc cca gtc agt gac act cca gat gaa ggg gat gag ccc atg cct      96
306 Ser Pro Pro Val Ser Asp Thr Pro Asp Glu Gly Asp Glu Pro Met Pro
307          20          25          30
309 gtc cct gag gac ctg tcc act acc tct gga gca cag cag aac tcc aag      144
310 Val Pro Glu Asp Leu Ser Thr Thr Ser Gly Ala Gln Gln Asn Ser Lys
311          35          40          45
313 agt gat cga ggc atg gcc agt aat gtt aaa gta gag act cag agt gat      192
314 Ser Asp Arg Gly Met Ala Ser Asn Val Lys Val Glu Thr Gln Ser Asp
315          50          55          60
317 gaa gag aat ggg cgt gcc tgt gaa atg aat ggg gaa gaa tgt gca gag      240
318 Glu Glu Asn Gly Arg Ala Cys Glu Met Asn Gly Glu Glu Cys Ala Glu
319 65          70          75          80
321 gat tta cga atg ctt gat gcc tcg gga gag aaa atg aat ggc tcc cac      288
322 Asp Leu Arg Met Leu Asp Ala Ser Gly Glu Lys Met Asn Gly Ser His
323          85          90          95
325 agg gac caa ggc agc tcg gct ttg tca gga gtt gga ggc att cga ctt      336
326 Arg Asp Gln Gly Ser Ser Ala Leu Ser Gly Val Gly Gly Ile Arg Leu
327          100          105          110

```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/755,830

DATE: 08/17/2001

TIME: 12:02:59

Input Set : A:\10287-067001.TXT

Output Set: N:\CRF3\08162001\I755830.raw

L:939 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:947 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:959 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:961 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:963 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:967 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:969 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:971 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:973 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:975 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:977 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:979 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:981 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:983 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:985 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:987 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:989 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:991 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:997 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:1256 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36
L:1258 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36